Inégalités : courbe de Lorenz, indice de Gini

Mise à jour : 04/03/2009

Ce texte présente deux outils fréquemment utilisés pour mesurer des inégalités (de revenu, de niveau d'études, etc.) entre personnes d'une même population ou entre populations distinctes : la courbe de Lorenz et l'indice de Gini

Introduction

Selon [1] page 55, au Ve siècle avant Jésus-Christ, le philosophe Platon prévenait les Athéniens :

> *« Il ne faut pas que certains citoyens souffrent de la pauvreté, tandis que d'autres sont riches, parce que ces deux états sont causes de dissensions. »*

Les inégalités entre personnes ou entre populations choquent notre sens de l'équité, qu'il s'agisse de revenus, de niveau d'études, d'accès aux soins, etc.

Dans ce qui suit, nous ne considérons que les inégalités définies par une grandeur mesurable, comme « le revenu disponible des ménages » ou « le nombre d'années d'études » ; nous excluons toute notion non quantifiable comme « la qualité de la vie ».

La première idée qui vient à l'esprit pour comparer les revenus des populations de deux pays consiste à utiliser une moyenne comme le PIB par habitant en tenant compte du coût de la vie, c'est-à-dire à parité de pouvoir d'achat (PPA). Mais lorsqu'il y a une forte différence entre ces PIB à PPA, comme c'est le cas entre ceux des Etats-Unis et de la France qui sont en 2008 dans un rapport voisin de 1,47, on ne sait pas si « les pauvres sont plus pauvres aux Etats-Unis qu'en France » : un revenu moyen comme le PIB à PPA n'indique pas la dispersion des revenus du pays.

On peut alors songer à caractériser la distribution des revenus d'une population en utilisant un second paramètre comme *l'écart absolu moyen*, qui est la moyenne des valeurs absolues des écarts à la moyenne :

> Soient n le nombre de personnes d'une population et r_1, $r_2 \ldots r_n$ les revenus de chacune. Le revenu moyen de la population est alors $r_m = (r_1 + r_2 + \ldots + r_n)/n$. L'écart entre le revenu de la personne de rang i et cette moyenne est, en valeur absolue, $e_i = |r_m - r_i|$ et l'écart absolu moyen est la moyenne des e_i, c'est-à-dire $(e_1 + e_2 + \ldots + e_n)/n$.

Mais la comparaison des écarts absolus moyens de deux populations ne permet pas de comparer correctement les inégalités résultant des dispersions des revenus de ces deux populations. En effet, un même écart absolu moyen de 2000€ représentait en 2008 moins de dispersion des revenus aux Etats-Unis, où le PIB à PPA moyen était de 48 000€, qu'en France où il était de 32 700€.

Nous allons voir ci-dessous une approche permettant de comparer correctement les dispersions de distributions statistiques, comme celles des revenus, des niveaux d'études, etc.

Courbe de Lorenz

Supposons connue la distribution des revenus d'une population de n personnes, c'est-à-dire les n nombres r_1, $r_2 \ldots r_n$.

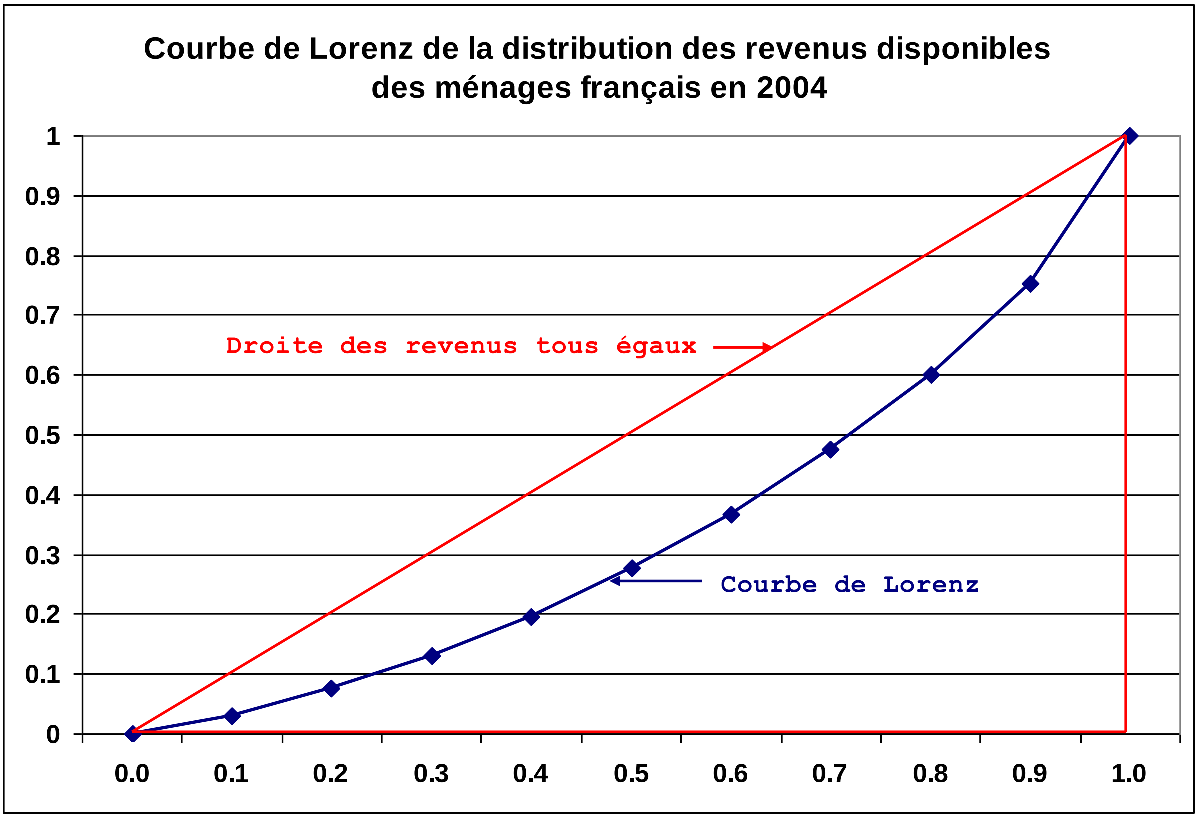
1. On calcule d'abord le revenu total de la population $M = r_1 + r_2 + \ldots + r_n$.
2. On range ensuite les n revenus individuels par ordre croissant et on les regroupe en classes de revenu croissant, par exemple comme ceci :
   - La première classe regroupe les 10 % de la population dont le revenu est le plus faible ; elle compte donc n/10 personnes.
   - La deuxième classe regroupe les 20 % de la population dont le revenu est le plus faible ; elle compte donc 2n/10 personnes et comprend toutes les personnes de la première classe.
   - La troisième classe regroupe les 30 % de la population dont le revenu est le plus faible ; elle compte donc 3n/10 personnes et comprend toutes les personnes des deux premières classes, etc.
   - La dixième et dernière classe comprend toute la population.

   La courbe de Lorenz relie la suite des points représentant la progression de la fonction "cumul des revenus d'une classe donnée" (y, en ordonnées) lorsque la variable "effectif de la classe" (x, en abscisses) augmente.

   La variable x a pour suite de valeurs les effectifs des classes successives, c'est-à-dire les pourcentages successifs de n : 10 %, 20 %, 30 %…100 %. La fonction y a pour valeurs la suite des cumuls de revenus des classes successives, cumuls exprimés en pourcentage du revenu total M de la population.

   On aura ainsi, par exemple (voir courbe ci-dessous) :
   - Une première classe de revenus groupant les x_1 = 10 % de la population dont les revenus sont les plus faibles et dont la somme y_1 des revenus représente y_1 = 3 % du revenu total M ;
   - Une deuxième classe regroupant les x_2 = 20 % de la population dont les revenus sont les plus faibles, classe qui comprend la précédente et dont la somme y_2 des revenus représente y_2 = 11 % du revenu total M ;
   - Une troisième classe regroupant les x_3 = 30 % de la population dont les revenus sont les plus faibles, une quatrième regroupant les premiers 40 %, etc. On arrivera ainsi à une neuvième classe regroupant les 90 % de la population dont les revenus sont les plus faibles, classe qui dispose de 79 % du revenu total M ; la dernière classe regroupe toute la population, qui dispose du revenu total M.
3. On relie alors les points successifs (x,y). Par convention, le premier point est (0,0) : les zéro personnes dont le revenu est nul ont pour revenu cumulé 0. Cette courbe est appelée *courbe de Lorenz* des revenus de la population considérée.

Exemple donné par [2] : Courbe de Lorenz du revenu disponible par unité de consommation [3] de la population française en 2004 (en bleu).



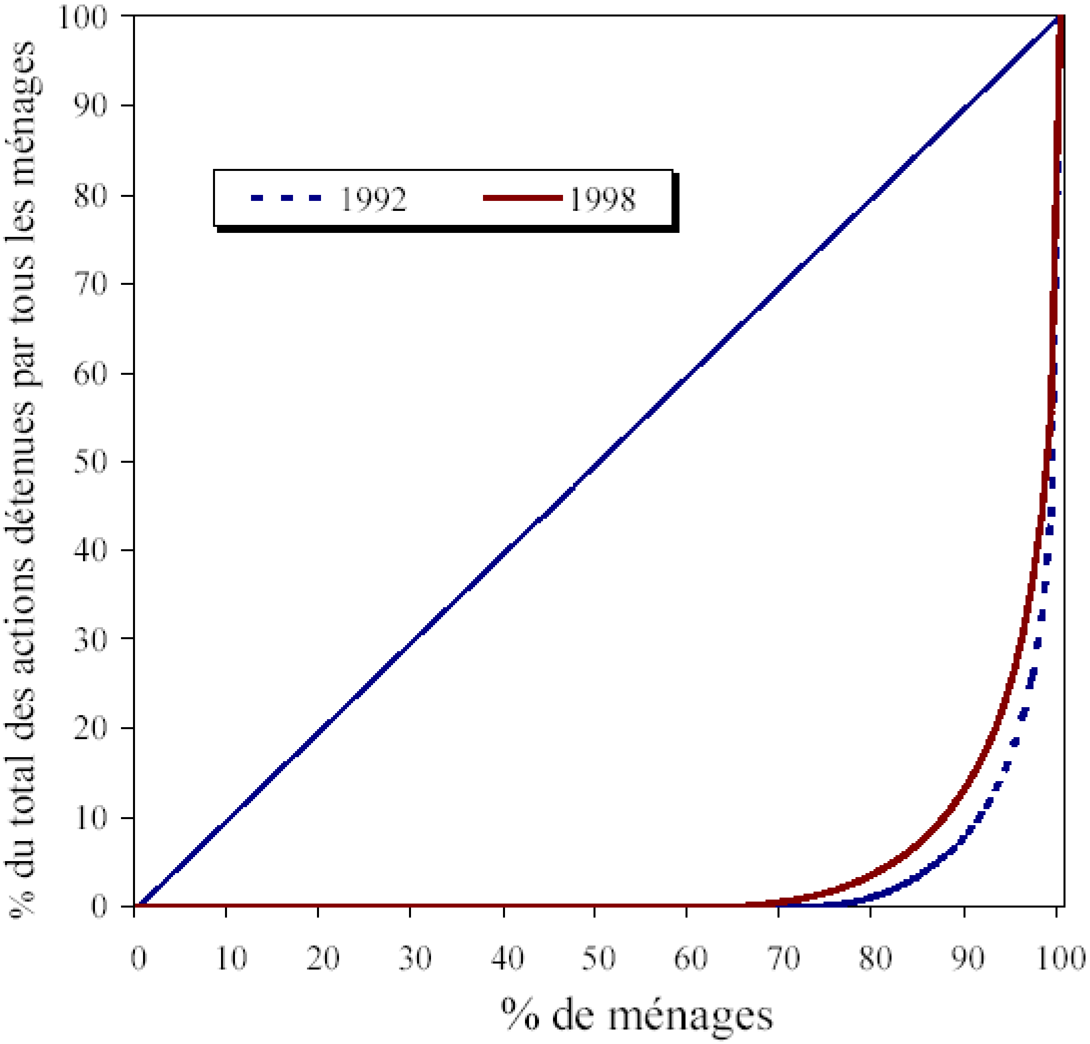
Courbe de Lorenz de la distribution des revenus disponibles
En abscisses : part dans la population totale - En ordonnées : part dans le revenu total

L'indice de Gini (défini ci-dessous) est égal au rapport : aire de la lentille entre la droite des revenus égaux et la courbe de Lorenz / aire du triangle rectangle rouge (qui vaut toujours 0.5). En 2004 il valait 0.268, après avoir constamment baissé depuis 1970.

Remarques sur cette courbe

- Cette courbe a une pente constamment croissante, car les revenus de la population ayant été rangés par ordre croissant, les classes successives ont une part du revenu total M qui croît de plus en plus vite.
- Si toute la population disposait du même revenu il n'y aurait qu'une seule classe, qui comprendrait 100 % des personnes et totaliserait 100 % du revenu M. La classe (vide) qui aurait 0 % du revenu M étant représentée par le point (0,0) la courbe de Lorenz serait alors le segment de droite joignant les points (0,0) et (100,100), c'est-à-dire la diagonale du rectangle (0,0) - (100,0) - (100,100) - (0,100). Cette diagonale est « la ligne d'égalité parfaite », inclinée à 45 degrés.
- Si au contraire l'immense majorité de la population avait un revenu très faible et un tout petit nombre de personnes un revenu très élevé, la courbe serait « très creuse » : à partir du point (0,0) elle longerait l'axe des x presque jusqu'au point (100,0) puis monterait brusquement jusqu'au point (100,100).

C'est le cas dans l'exemple ci-dessous, représentant la distribution du patrimoine en actions en France en 1992 et 1998 d'après [4] :

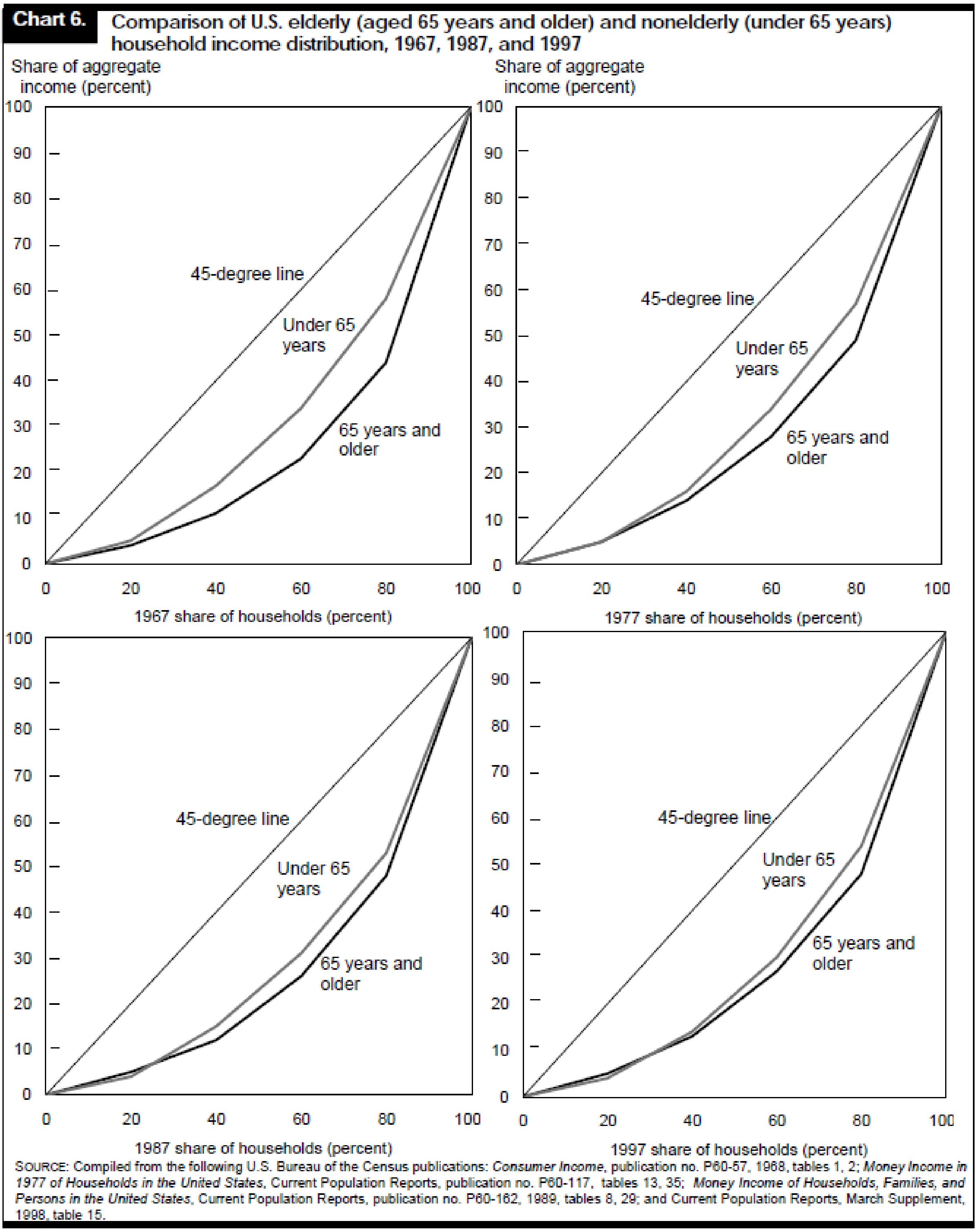


On voit qu'un petit nombre de ménages détiennent l'immense majorité des actions, et que cette situation d'inégalité s'est un peu améliorée de 1992 à 1998.

Remarque importante

Il n'y a pas de raison de considérer des classes de population de même effectif, 10 % jusqu'ici : les effectifs peuvent être quelconques. Et comme il est intéressant de préciser la courbe au voisinage de son dernier point, (100,100), on peut prendre par exemple comme dernières classes les valeurs de x correspondant à 90 %, 95 %, 99 % et 99.9 %.

Autre exemple, issu de [5] :



Courbes de Lorenz comparant les distributions des revenus des personnes de moins de 65 ans et des plus de 65 ans en 1967, 1977, 1987 et 1997 aux Etats-Unis

La courbe de Lorenz a donc l'avantage de représenter une population par une fonction y(x) telle que y et x soient tous deux des pourcentages, donc des nombres sans dimension et convenant à n'importe quelle population, quels que soient la moyenne et l'écart absolu moyen de sa distribution.

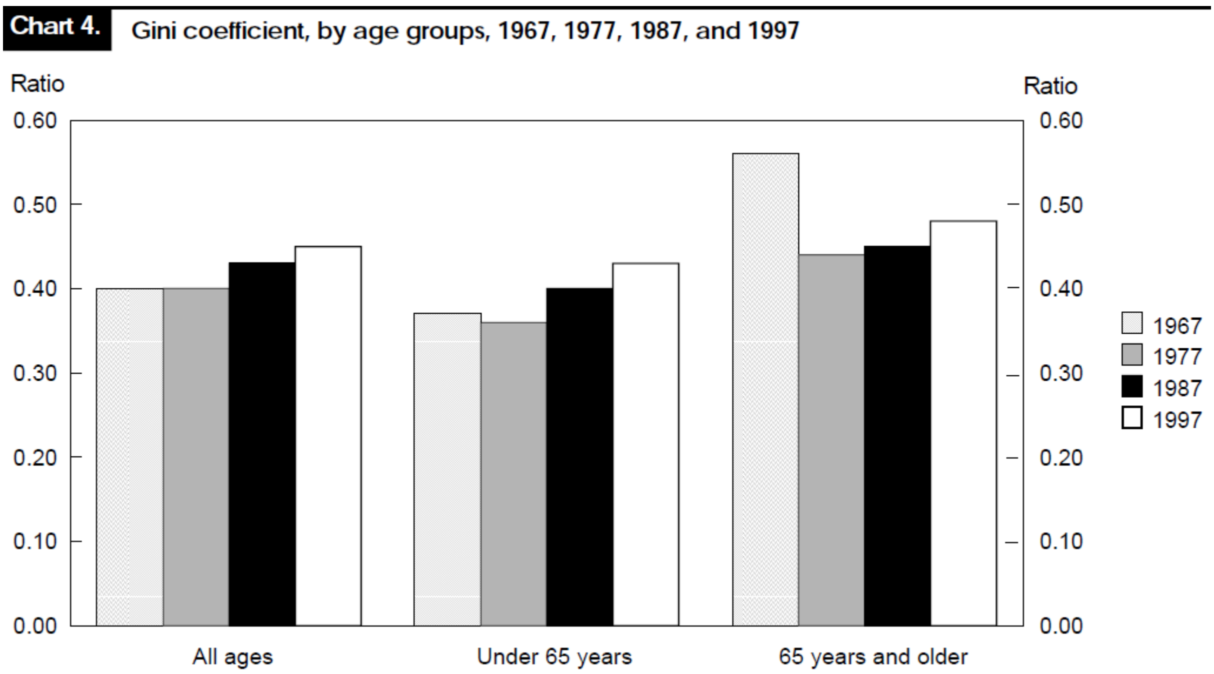
Indice et coefficient de Gini

Considérons la surface en forme de lentille comprise entre la diagonale d'égalité parfaite et la courbe de Lorenz, et soit S son aire. Plus les inégalités sont importantes dans la population considérée, plus la courbe de Lorenz est "creuse" et proche de l'axe des x, plus le rapport G entre l'aire S et l'aire du triangle rectangle sous la diagonale (0,0) - (100,0) - (100,100) est important.

Le rapport G, appelé *indice de Gini*, est donc une mesure des inégalités. Sa valeur est comprise entre 0 et 1. Plus G est grand, plus les inégalités entre personnes de la population considérée sont importantes. *On peut donc comparer les inégalités des distributions de deux populations en comparant leurs indices de Gini G*.

On exprime parfois l'indice de Gini en pour-cent en parlant de **coefficient de Gini**. C'est ainsi qu'un indice de Gini de 0.75 équivaut à un coefficient de Gini de 75.

Voici un graphique complétant le graphique de [5] ci-dessus par la variation des indices de Gini, et y ajoutant ceux de la population américaine toute entière :



Coefficients de Gini des distributions des revenus des Américains : on voit que les inégalités ont tendance à croître depuis 1977 dans tous les groupes d'âges

En 2007, l'indice de Gini de la distribution des revenus des Américains était de 45 contre 28 en France en 2005. Comme cet indice ne varie que lentement avec le temps (cela se voit sur le graphique ci-dessus), on peut affirmer que *les inégalités de revenus sont bien plus faibles en France qu'aux Etats-Unis*.

> Si on considère pour la France *le niveau de vie* (mesuré par le revenu disponible) au lieu *des revenus bruts*, le coefficient de Gini 2002 est 26.7 au lieu de 32.7 [7]. C'est là l'effet redistributif des transferts sociaux et des impôts, effet considérable en France.

Exemple issu de [1] page 55 : comparaison des inégalités de revenus à l'intérieur d'un certain nombre de pays ou de groupes de pays :

Pays ou groupe de pays	Coefficient de Gini du PIB par habitant à parité de pouvoir d'achat
Suède	25.0
France	32.7
Fédération de Russie	31.0
Egypte	34.4
Royaume-Uni	36.0
Pays à revenus élevés de l'OCDE	36.8
Chine	44.7
Etats-Unis	40.8
Brésil	59.3
Monde entier	67.0
Afrique subsaharienne	72.2

Coefficients de Gini de la distribution des revenus connus en 2005

On voit que la Suède est (dans cette comparaison) le pays où les revenus bruts sont les moins inégaux et que l'Afrique subsaharienne est la région où ils sont les plus inégaux.

En résumé, l'indice de Gini est le rapport entre l'aire comprise entre la "diagonale des revenus égaux" et la courbe de Lorenz, et l'aire du triangle sous la diagonale : plus ce rapport est élevé, plus les inégalités sont fortes.

Cette approche permet de comparer les inégalités de pays dont les revenus sont assez différents, car dans la fonction de Lorenz y(x), la variable x et la fonction y sont tous deux de simples nombres sans dimension, compris entre 0 et 1 (ou 0 et 100).

Daniel MARTIN

Références

[1] RAPPORT MONDIAL SUR LE DÉVELOPPEMENT HUMAIN 2005 des Nations unies http://hdr.undp.org/en/media/HDR05_complete.pdf

[2] INSEE - Distribution des revenus disponibles des ménages 1996-2004 http://81.255.68.41/fr/ffc/docs_ffc/revenus_pauvrete/TW-revdisp-distribution.xls

[3] La notion d'unité de consommation (uc) permet de comparer les revenus de ménages de taille différente : on compte une uc pour le premier adulte, 0,5 uc pour chaque autre membre du ménage de 14 ans et plus, et 0,3 uc pour chaque enfant de moins de 14 ans.

[4] Document du Sénat http://www.senat.fr/rap/r02-367/r02-3676.html téléchargé le 04/10/2005

[5] Monthly Labor Review, November 2000 http://www.bls.gov/opub/mlr/2000/11/art2full.pdf

[6] INSEE - France portrait social 2005 (Fiches thématiques) http://www.insee.fr/fr/ppp/sommaire/FPORSOC05.PDF
En France, l'indice de Gini du niveau de vie est inférieur à celui qu'on trouve pour les revenus, parce qu'il y a un fort effet de redistribution dû aux transferts sociaux et aux impôts.

Retour page d'accueil